

---

# ENGLISH FOR LINGUISTICS AND MULTIMODAL PEER-ASSESSMENT AT UNIVERSITY POSTGRADUATE LEVEL

Maria Grazia Sindoni  
(*University of Messina, Italy*)

## Abstract

Is English for Linguistics (EL) a domain of interest for EAP? Is the metalanguage for linguistics (e.g. lexical precision, semantic and pragmatic appropriateness) sufficiently taught at university level? Which strategies are most appropriate when developing presentation skills with regard to language competence in the field of linguistics?

This paper sets out to address these questions, adopting the viewpoint that competence in EL is probably taken for granted at university level and less researched than it should be. Strategies to encourage the development of this particular metalanguage, with reference to specific lexical items and semantic areas, are investigated in peer-assessment procedures, which would seem to be particularly effective at postgraduate level when integrating syllabus content and language skills to negotiate and reflect critically on this aspect of EAP.

Despite general agreement over the usefulness and impact of peer-assisted educational strategies (Topping 1988; Falchikov 2001), there is a striking lack of experimentation on peer assessment, especially when it comes to formal recognition and inclusion in university syllabuses within EAP practice. The rationale of this paper builds on a pilot project carried out at the University of Messina (Italy) in 2010, in a course of English Linguistics for postgraduate students in Foreign Languages and Literatures in which systemic-functional and cross-cultural socio-semiotic approaches to multimodal studies (Baldry & Thibault 2006; Kress & van Leeuwen 2006) were the major focus of analysis. Part of the course consisted in the development of individual projects, assessed both by the teacher and their peers with the ultimate goal of developing reflective, linguistic, metalinguistic and presentation skills.

Related issues are discussed, such as students' development of assessment grids, the integration of contents and metalanguage, and the consistency between peer and teacher evaluations. This approach helps expand students' language autonomy in articulating evaluative decisions and priorities regarding their own and their peers' learning outcomes. The mastery of a specialized language is targeted both as regards discussing syllabus contents and as regards expanding expertise in the field of linguistics.

## 1. English for Linguistics and peer assessment: the rationale for an EAP course

Starting from the consideration that *English for Linguistics* (EL) may be considered as a specific area of English, including a highly specialized lexicon and other preferred structures for a specific target audience (e.g. university students of Arts, Humanities and Education and prospective teachers of all levels in English as a Foreign Language), it may well be argued that the gap between experimentation and research in EL needs

to be filled. Furthermore, if we take into account the crucial importance of teaching and assessing future teachers in the field of EFL, EL becomes a significant, though neglected, area of investigation. The far-ranging impact of communicative approaches in teaching and learning foreign languages has placed more emphasis on strategies for effective and fluent communication, putting aside such thorny areas as formal grammar descriptions, with their specific and sometimes complex labelling. Such formal descriptions are, more often than not, considered as the Achilles' heel for both teachers and students, but they need to be taught and learnt especially in the context of university training programmes for future teachers of EFL.

A second fundamental consideration in the present discussion is the observation that university students in the areas of Arts, Humanities and Education often lack sufficient training in the fundamental area of testing and assessment. In Italy, where the experiment discussed in this paper has taken place, prospective teachers of EFL lack specific training in both teaching and assessing skills. Postgraduate degree courses often fail to provide practical training programmes. Moreover, courses in linguistics usually focus on theoretical issues, requiring students to develop theoretical expertise to be used typically for an oral exam, a written paper or a final dissertation. EL will be thus considered as an overall domain of investigation for the development of a specific metalanguage used to teach, but also to assess. It is, as such, mainly targeted at encouraging students to recognize the value and usefulness of using this metalanguage effectively. This paper is an attempt to put together the two strands of EL and peer assessment within a vocational context, also considering how to develop presentation skills.

In the last two decades considerable academic and educational effort has been devoted to the exploration of a wealth of learning and assessment orientations and procedures which have been steadily shifting from a rigid teacher-led perspective to a student-centred approach (Alderson & North 1991; Falchikov & Goldfinch 2000; Falchikov 2001; Boud & Falchikov 2007), also in the field of ESP (Hutchinson & Waters 1987; Douglas 2000).

Different educational theories converge on the general consideration that forms of self learning, self assessment, peer monitoring, peer pairing and peer assessment greatly enhance the experience of learning, improving conditions, strategies and outcomes. Many approaches hold that peers can boost the conceptual, emotive, intellectual, cognitive and metacognitive development of their partners, encouraging a more student-centred classroom (Stiggins 1994). Methods for peer learning range from cooperative learning (Johnson & Johnson 1987) to collaborative learning (Brown & Campione 1994) and peer tutoring (Cohen, Kulik & Kulik 1982; Greenwood 1997). These methods vary in the application of peer learning, but they generally agree about its usefulness and positive backwash in educational achievements. O'Donnell & Topping's (1998: 259) early claim that research literature on the use of peers for assessment was "quite sparse" proves still valid in relatively recent studies (Liu & Carless 2006; Callahan 2007; Frankland 2007) and also very recent studies (Kaufman & Schunn 2011; Jin 2012; McConlogue 2012). Why, if peer learning is so commonly held to be effective and positive, is there a striking lack of experimentation regarding

peer assessment, especially when it comes to its formal inclusion in the syllabus and curriculum and in particular in EAP contexts?

A central issue concerning the partial lack of systematic use of peer assessment in EAP contexts lies in the distinction between formative and summative peer assessment. The former deals with the process of learning and may be better defined as peer monitoring, that is, helping the partner/s with critical feedback and providing support in terms of advice and *in itinere* group work response and evaluations. In other words, formative peer assessment is more concerned with the process and gives the opportunity to revise the product to be assessed before handing it in, and this is particularly relevant within EAP evaluation. Summative peer assessment is instead concerned with the final outcome, i.e. the product of learning after a period of instruction. Summative peer assessment is typically designed as a way to grade peer work (e.g. essays, presentations) and is connected with achievement. Whereas summative peer assessment is not based on purely “objective” marking criteria (e.g. univocal answer cloze questions or correcting grids), students may feel uneasy about their own grading or suspicious about their peers’ grading. Moreover, teachers may have more than one reason to fear peer assessment, being wary of their students’ lack of expertise, training and of other more covert issues, such as giving away a part of their institutional power to students. A partial reversal of institutionalized roles is perhaps what makes teachers (and curriculum planners) so resistant to formalized summative peer assessment. Educational planners in general are cagey about formalizing peer assessment, whereas self and peer learning constitute a common ground of investigation and experimentation, for example in language planning, and have been thoroughly institutionalized at European level (see the European Language Portfolio).

Students can in effect be controversial assessors. If placed in the role of evaluators, reliability and validity may be at risk. Pond *et al.* (1995) listed many controversial issues, such as *friendship grading* (i.e. students assigning high grades to peers because of friendship), *collusive grading* (i.e. lack of differentiation between peers, especially frequent with high stake assessment), *decibel grading* (i.e. students assigning the highest grades to the most active peers). In the experiment discussed in this paper, other controversial factors were the very competitive environment where peer assessment was implemented, non-existent experience as regards peer assessment and a low degree of familiarity with EL outside the oral exam context. However, peer assessment within EAP courses is experimented less often than it should be, even in undergraduate and postgraduate university contexts, where some of the possible problems could be countered by the students’ (hopefully) highly developed critical skills.

However, experimentation needs to be carried out further, especially at university where students are required to improve their negotiating and evaluating skills in EL, particularly in the Humanities, where a significant number of students need to develop vocational skills for future teaching careers. In Italy, many teachers complain about the poor quality and virtual lack of practical teacher training at postgraduate level. A solution is more and better teacher training, of which peer-assessment is a small part. However, experimentation and research into peer assessment within EL is needed to

enhance university students' learning experience and to equip them with practical tools to become assessors in (future) real life educational contexts.

The quality of design is deemed essential in any curriculum or syllabus where self or peer learning or assessment is involved, including ESP courses (Dudley-Evans & St. John 1998; Boud & Falchikov 2007). This is particularly true about peer assessment in EAP contexts, as students need to be clearly instructed in what they are supposed to do, given transparent and unambiguous criteria to assess their peers, trained to peer monitoring, and given both positive and critical feedback regarding their peers' work. Development of assessment grids and negotiation of assessment values has proved to be a crucial factor in many case studies (Liu & Carless 2006), and this is particularly true in the case of ESP testing and assessment (Dudley-Evans & St. John 1998).

Another crucial factor in designing the course is how to measure the success of the experiment. The usual measurement of success in similar experiments was the degree of agreement between teacher and student ratings (Falchikov 2001). However, in keeping with what Falchikov (*ibid.*: 272) herself claims as regards success in peer assessment, "agreement between student and teacher marks may not be the most important aspect of successful self- or peer assessment. Real success should follow from the enhancement of student learning that results from participation in the process". Measurement of success cannot be exclusively equated with the agreement of grading between students and teacher for a number of reasons, such as the consideration that no grading may be believed as a pure or neutral benchmark. Teachers' grading is more subjective and evanescent than we, as teachers, are willing to admit, and especially so when it comes to marking via complex and non-univocal criteria (e.g. essays, presentations, oral exams). Another reason for the need to expand our notion of success in peer assessment experimentation lies in the ideological consideration that empowering students and fostering their reflective skills with regard to the complex arena of EAP assessment is a task well worth undertaking.

## 2. Multimodal peer assessment within EAP testing

Peer assessment is a neutral label, since it may be defined as a method that needs to be complemented by a theory of ESP assessment. Assessment, in turn, is part of an EAP testing process. Dudley-Evans & St. John (1998) claim that assessment does not stand alone, but is part of a process where needs analysis, course and syllabus design, teaching/learning and evaluation interact and affect one another. In principle, any ESP test can be defined as a performance test, assessing the skills needed to perform successfully in the interplay between language knowledge and specific purpose content knowledge (Tratnik 2008). Douglas (2000: 10) argues that ESP tests are "contrived language use events" in which the test takers' specific purpose language skills and knowledge of the specialist field are measured. ESP tests are related in content, themes and topics to particular fields of studies and, as such, measure the degree of development of language specificity skills. Specific lexical, semantic, syntactic, pragmatic and cross-cultural features need to be taken into account in assessing EAP.

From this standpoint, lexical precision is of special relevance in EAP testing and assessment within EL discourse practices. As ESP testing is generally constructed around the demands of specific workplaces and language situations, tests should include tasks that reflect those needed by ESP test takers. In this specific case, students needed to develop and consequently be able to show awareness and competence in their use of the metalanguage to teach EL and to assess in and through EL, foregrounding presentation skills. The definition of EL tries to capture the complexity of the language used by teachers of EFL: it is something more than the mere knowledge of grammar, vocabulary, etc. as it encompasses a broader approach to language and communication, taking into account different theories of grammar, in this case systemic-functional approaches to grammar and the semiosis of cross-cultural communication from a multimodal standpoint.

### 2.1. *Course design, planning and implementation*

These observations led to the development of an EAP peer assessment project at the University of Messina, Faculty of Arts and Humanities, in a postgraduate course on English Linguistics. Preliminarily a survey was developed and administered to students to gather information for a needs analysis. When prompted about their needs and *desiderata*, students complained about the alleged lack of practical training with regard to teaching and assessing EFL at university at both first and second level degree courses.

The course was accordingly designed to provide theoretical tools and to put students in practical hypothetical future target situations, e.g. teaching EFL with multimodal texts. My own experience of testing and assessment (Sindoni, Cambria & Stagno d'Alcontres 2007; Sindoni & Rizzo 2008; Sindoni 2009; Sindoni & Rizzo 2009; Sindoni & Cambria 2010) revealed the need to constantly re-think and re-engage with assumptions, ideas, even theoretical frameworks, if we aspire to keep high stake tests in tune with a constantly changing learning environment and students' vocational needs. We should not content ourselves with the bare basics of testing and assessment. There are also ethical implications (Sindoni & Cambria 2010) regarding the impact that tests and exams play on individuals' lives and on society in general, not to mention other preliminary factors, such as what to measure and what to consider as evidence of "learning" in an EAP course specifically dealing with EL.

Quantitative and qualitative methods have been developed within testing theories, serving the purpose of evaluating test rationale and formats via different criteria, such as the classic notions of validity, reliability (cf. Hughes 2003 [1989]) and examining *a priori* and *a posteriori* validity evidence (Weir 2005). However, experimentation on peer assessment in an EAP context may shed light on critical issues such as, to put it simply, the embedding of form and content, that is, the incorporation of specialized language with course contents. A crucial factor in encouraging and developing the required expertise in students is likely to be found in de-mythicizing the teacher's role, which may be, on the one hand, uncritically internalized or, on the other hand, resented or rejected by university students. Both positions imply a somewhat passive acceptance of pre-established teaching roles and need to be contested in the positive dynamics of teacher-student negotiation.

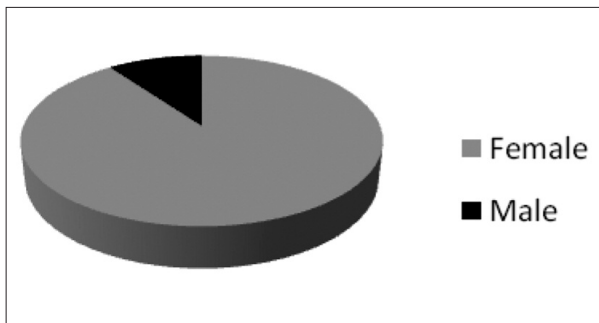
Moreover, students take for granted that they are completely passive when it comes to evaluating the results of their learning. Peer assessment experimentation was an attempt to counter all these problems. In addition to this, the relationship between content and language is strengthened and made more relevant and significant within classroom practices accordingly.

As the course was on systemic-functional grammar and multimodal studies, as developed within the theoretical frameworks by Baldry & Thibault (2006) and Kress & van Leeuwen (2001, 2006), there was a need to put a multimodal lens in front of the students' eyes, also focusing on cross-cultural hybridization made possible by the intersection of such theoretical approaches. The idea was then to integrate the content of the course with methods echoing the complex nature of the texts analysed (e.g. music videos, adverts, written texts, website homepages, trailers) highlighting the importance of the *contexts* where these texts originated, in compliance with notions of "context of culture" and "context of situation" (Malinowski 1923; Halliday 1978; Halliday & Matthiessen 2004) and considering texts as specific examples stemming from specific cultural contexts. To counter potential essentialist views on texts and genres, devoid of cultural implications, students were encouraged to critically think of text or genre differences across different cultures and report on their reflections in their presentations. Furthermore, crucial semantic areas were targeted during the course and students were prompted to pay attention to their peers' competent use of those areas during assessing sessions. These areas included specialized lexicon drawing from systemic-functional grammar labels and multimodal definitions. Students were encouraged to use precise definitions and accurate systemic labels in order to get the message across to the audience. Moreover, they were also encouraged to embed the notion of register, presented as a kernel notion of the course, into their presentations and to use what they knew about register and context to use the appropriate register while delivering presentations.

Students were enrolled in the university postgraduate degree in "Foreign Languages and Literatures", and the course on "English linguistics" was compulsory for first and second year students. Table 1 provides details with regard to the course. However, they had the opportunity to enrol in the experiment on a voluntary basis. The project involved the development of a presentation on a topic which revolved around the course syllabus, namely multimodal text analysis, preferably exploring contemporary text genres and comparing them across different English-speaking cultures. Students were informed that the project was on peer assessment, and that meant assessing and being assessed for their class presentations. Twelve students participated in the project out of the 60 attending. They were aware that they were also responsible for a part of their peers' final and formal assessment, with their average marks accounting for 30% of their peers' final mark out of a maximum of 30. But what about the rest of the class? Since it was unlikely that all students would be willing to participate, a partial solution to the potential risk of marginalizing those who would not join the project was found in their involvement in the development of the assessment grid. In other words, all students participated in the assessing procedures, but only those who did the presentations had a formal role in assessment, i.e. their average mark was computed in their peers' final course mark.

Number of students	80
Number of attending students	60
Age	22-26
Participation in the project	20% of attending students
Participation in the Facebook discussion group "Our old friend Halliday"	100% of attending students
Agreement between teacher and students grading	84%
Course duration	30 hours with teacher plus 50 hours of English language skills with language tutor
Course	English Linguistics
Language level	Advanced to proficient (C1-C2)
University level	Postgraduate degree in Foreign Languages and Literatures

*Table 1.* Background information on the course



*Graph 1.* Gender of participants

Since students had no previous experience either in peer learning or assessment, ample class discussion was devoted to debating the implications of such an experiment. Students expressed doubts and felt challenged by the peer assessment undertaking. Their main perplexity concerned their lack of training and the underestimation of their ability to get to

grips with anything that involved assessment. Preliminary work was then focused on the following areas:

- explaining the syllabus design and the ideological and theoretical rationale underlying it
- building trust and eroding the culture-bound competitive learning environment
- setting clear purposes, e.g. providing quality feedback to be acted upon, empowering and challenging oneself, assessing and being assessed on collectively negotiated criteria
- developing assessment grid in and outside class with the help of teacher and language tutor
- creating a Facebook discussion group “Our old friend Halliday”, where all the issues could be debated and collaboratively addressed, with constant teacher monitoring.

The development of specific competence in EL was thus fostered via the identification of the following areas:

- lexical: encouraging students to use specific labels from systemic-functional grammar, such as “experiential, interpersonal and textual metafunctions”, “transitivity patterns”, “material, mental, relational, attributive, existential processes”, “theme”, “rheme”, “modality”, “appraisal”, “field”, “tenor”, “mode”.

- semantic: helping students focus on specific semantic areas in their presentations related to the course contents, such as functional grammar and multimodality;
- pragmatic: highlighting how communication is much more successful when a specific metalanguage is used appropriately during class presentations;
- cross-cultural: comparing how different texts or genres are constructed in different English-speaking countries.

As each presentation dealt with a different text and a related text analysis, students soon learned to be flexible and highly responsive to each of their peers' performance in customized terms. The assessing grid was rather flat and did not allow specific observations; this role was thus played by the comment section which was used extensively to express specific concerns about lexical, semantic and pragmatic issues. What was previously perceived as a rather passive reception of teachers' contents was then appreciated as an active construal of texts and peer performances according to 87% of students.

## 2.2. Discussion of results

As research literature attests (Falchikov 2001), students generally report that they are facilitated when involved in the discussion and elaboration of the assessment criteria. Students in the project discussed here were no exception. All students engaged in the design of the assessment grid. Since the course was focused on systemic-functional linguistics and multimodal studies, they had been previously trained in expanding their notions regarding language and communication. For example, among the selected and to-be-assessed criteria, special attention was placed on visual aspects (i.e. in their presentation) and communicative skills which, they were well aware, are not exclusively based on verbal abilities or in the mastery of the foreign language (i.e. English), but rest on a wide range of resources, such as the ability to involve the audience. Students were comfortable with traditional and "measurable" criteria of assessment, for example those related to traditional language skills (e.g. *Is the presenter using a fluent, correct, appropriate language?*) or covering the proposed topic in a more or less exhaustive way (e.g. *Is the presenter covering the topic with enough examples? Is the point clear? Is the presenter using both practical examples and theory to pinpoint her/his discussion? Is the presenter satisfactorily answering the audience's questions?*). Points related to language skills, coverage and general "knowledge" of the presented topic were both maximally valued and easily recognized by students.

Less measurable criteria were not only hard to identify, but also hard to assess. For example, the general ability to deliver a presentation, which nonetheless was perceived as fundamental in the overall appreciation of it, proved hard to pin down and identify. The same may be said about notions such as register, which risked being too evanescent to be measured.

As the course progressed and core concepts were introduced (e.g. language as a system, meaning potential and behaviour potential, the grammar of visual communication, the meaning compression principle), students soon became aware that language and communication are much more complex than they had thought. The multimodal study of a wide range of texts required fine-grained theories and ensuing descriptive models. A multimodal analysis was thus complete and meaningful only if



it implied the use of a multimodal lens with which to read textual phenomena and if it was studied and discussed “multimodally”. The link between syllabus content, methods of lectures and assessment left students at a loss at the beginning. They were not used to establishing a connection between what they learned, how they learned it and how and about which acquired skills they were assessed. For example, they found it hard to identify the qualities of a presentation which are not optional but an integral part of communication and, as such, come into play in assessment. However, they soon started pondering about “delivery” resources, which are communicative and produce meanings as much as verbal communication. The identified “delivery” resources included gaze (e.g. eye-contact with audience), body language in the broadest sense (e.g. voice control, audibility), management of space and effective use of visual resources (e.g. *is the presentation readable?*). Sometimes very good presentations are difficult to follow due to purely visual factors, e. g. non-readability, bad colour matching, or due to a wrong combination of visual and verbal (excess of writing per slides, tautological/pedantic or useless repetition by speech of what is already written in the slide). Students were less trained in recognizing whole ranges of basic communicative strategies which, if poorly used, hinder successful communication. As regards lexical, semantic and pragmatic precision, they started thinking about the importance of precision when tackling linguistics, which requires clear, unambiguous and appropriate use of the metalanguage which was at the core of the course. Integration of content with language was addressed and discussed both during the development of the assessment grid and during peer assessment sessions.

In keeping with the aims of the course, students were encouraged to pay special attention to the integration of verbal and non-verbal modes, which are especially relevant when analysing cultural features. Students are usually much better trained in interpreting verbal (i.e. written or oral) texts: they attend classes, lectures and produce written and oral texts for assessment purposes. They very rarely or never assess their peers and are not frequently trained in identifying which non-verbal modes come into play in contemporary communicative events. However, after an initial trial period at the beginning of the course, they were eager to participate in such a peer assessment experiment. The cross-cultural dimension surfaces in the selected sample titles of presentations listed below:

- Gender and sexuality in ethnic-biased advertising
- Visual grammar and the body. Karate between East and West
- Different trailers in different cultures: *Alice in Wonderland*
- A multimodal text analysis of Harvard and Cambridge homepages
- How do we “read” colours? Semiotic “colour-readings” across cultures
- Theme and rheme in Eastern and Western cultures

Students elaborated the assessment grid after eight hours of class and group discussion. Quality feedback was provided and the grid was constantly revised to keep up with the group’s reflections and feedback after mock assessing sessions. Two cohorts of students tested the same assessing grid in separate sessions, whose final version is shown in Appendix 1.

Students found it very complex to squeeze into a single assessing grid the whole plethora of criteria they had identified during their mock assessing sessions. The final

grid includes three macro-categories with three sub-criteria each, for a total of nine different marks to be awarded during each presentation, ranging from most to least satisfactory marks (i.e. A-D). Both cohorts of students (made up of 25-30 students each) claimed that it had been very complex to reduce categories and sub-categories (which formerly amounted to more than 20) and that “assessing is very hard, especially paying attention to different things at the same time!”, as one student wrote in a comment posted in the discussion group.

Agreement of grading between students and myself amounted to 84% at the end of the scheduled presentations. The bottom part of the grid, which I thought at the beginning would have been left empty in the majority of cases, due to the students’ admitted difficulty in “paying attention to many things at the same time”, was the most successful part. They provided quality feedback, giving detailed and expert advice on a number of different aspects related to their peer performances, noting with striking accuracy a large quantity of details, such as how to improve presentational skills, manage time, be more effective and successful communicators and also specific comments related to lexical and semantic targeted areas.

What happened outside the class is also worth briefly reporting here. They were prompted to create a Facebook discussion group, which they called “Our old friend Halliday”, a tool to monitor their activities *in itinere*, check their fears, doubts and enthusiasm about their progress. The discussion group was teacher-monitored and qualitative analysis of posted comments allowed interesting insights into the students’ learning process from their own point of view.

Gender-related differences were evident in posts, but the small sample does not really permit generalization, even though my data agree with what has emerged from research literature (Hutchinson & Waters 1987; Topping 1988; Falchikov 2001). From a manual lexical survey on all (232) posted comments, it emerged that girls felt more insecure *before* the peer assessment procedures and felt threatened (cf. “fear”, “anxiety”, “scared”, “mistake”), while boys (who, incidentally, *all* joined the experiment) showed more confidence and buoyancy (“great”, “interesting”, “power”). However, when it came to discussing final results, the boys were eager to show their penchant for competitive verbal behaviour (e.g. “disappointment”, “offended”, “criticisms”, “anger”), while the girls were much more conciliatory, devoting a considerable amount of posts to appeasing conflict (e.g. “wonderful”, “best”, “happy”, “excellent”, “accept”). Competitive verbal behaviour was also evident in the number of posts devoted to debating the fairness of peer assessment results (about one-third of the total), for example challenging some aspects of the experiment design (e.g. alleged unfairness about the scheduling of presentations, as the first “icebreakers” were considered to be at a disadvantage compared to the subsequent presenters). These considerations also point to the fact that *any* experiment in EAP teaching/learning needs to take into full account cultural dimensions that may significantly alter course planning, design, implementation and evaluation.

From an informal survey carried out at the end of the experiment, 78% of students reported beneficial results, showing an overall agreement over the positive effect of having the responsibility of assessment. Among the side-effects of their increased sense of responsibility towards their peers, they claimed they had been more focused on

course contents and peer presentations and more active in the evaluation of their own and their partners' work.

### 3. Conclusions

As is evident from the previous discussion, this experiment is not focused on quantitative data. It is not easy to pin down how students assessed their peers in practical terms, as each presentation brought out different issues and different ways of integrating language use and content. As learning is more a process than a product, the development of the assessing grid has been the central focus of interest of this study. Sections devoted to the specific use of EL were discussed by students in class and in digital environments (i.e. Facebook discussion group). They developed awareness on the specific competence in EL, here considered as both a metalanguage and as an umbrella skill to teach and assess in EL and through EL. This double aim was reached via the identification of target areas (i.e. lexical, semantic, pragmatic) and the appreciation of how an inappropriate use of language and metalanguage may be detrimental to successful communication in their future target context, i.e. a class.

Students analysed a number of different text genres in their presentations, such as *Youtube* music videos, website homepages, visual grammars (e.g. karate visual manuals), children's educational videos and film trailers. Akin to the complexity of contemporary digital and non-digital text genres, they practised an approach which integrated content, methods and assessment procedures, refining their analyses with the help of cross-cultural reflections. Class discussion emphasized the ideological implications of such an approach, which broadens mainstream notions of language and communication and maximizes students' interest in providing critical prompts, elaborating "signs of learning" (Jewitt & Kress 2003) and turning them into signs of personal growth which is socially negotiated, culturally moulded and never acquired in isolation. However, it needs to be remarked that what was successful for a group of motivated second-level students is not necessarily deemed to work in another context.

The shift from authority and authorship to peer irradiating knowledge is central in our contemporary age and is instantiated in everyday systems of producing, exchanging and distributing information. Despite the expertise in the taught field, the teacher may and should be contested by appropriately trained students, whose critical approach to any subject needs to be encouraged via daring educational choices. The concept of the "wisdom of crowds", as developed by Surowiecki (2004), may also be applied to the assessing domain, even though it may destabilize our notions of authority in the educational domain and traditional systems of handing down knowledge and teaching skills and abilities in specialized contexts. The bottom line of this experiment is that accepting that our students may express wiser opinions (formulate judgements and possibly evaluations) than ours nonetheless undermines well-trodden certainties. However, a corresponding shift from notions of reliability in testing to the more blurred and challenging pedagogic assumptions underlying the peer assessment realm within EAP contexts is well worth exploring across the fruitful but conflicting lines which this paper has attempted to outline.

## Appendix 1

UNIVERSITA' DEGLI STUDI DI MESSINA  
 FACOLTÀ DI LETTERE E FILOSOFIA  
 Corso di Laurea magistrale in Lingue e Letterature Straniere  
 Lingua Inglese: Prof. Sindoni

### PEER ASSESSMENT

Name:

Date:

1 LANGUAGE SKILLS					
1.1	<b>Communication</b> (is the presenter using a correct and fluent language? - e.g. correct grammar and understandable pronunciation)	A	B	C	D
1.2	<b>Range of vocabulary</b> (is the presenter using both general and technical terms and a good range of specific lexical items?)	A	B	C	D
1.3	<b>Register</b> (is the language appropriate to the context?)	A	B	C	D
2 TOPIC					
2.1	<b>Coverage</b> (is the presenter covering the topic satisfactorily? Is the general point clear?)	A	B	C	D
2.2	<b>Informativity</b> (is the presenter providing enough examples? Is s/he linking theory and practice effectively?)	A	B	C	D
2.3	<b>Structure</b> (is the presentation well-organized? – e.g. title, introduction with explicit aims and methodology, data presentation, analysis and interpretation)	A	B	C	D
3 FEEDBACK					
3.1	<b>Interaction</b> (is the presenter involving the audience with questions? Is he/she satisfactorily answering the audience's questions?)	A	B	C	D
3.2	<b>Body Language</b> (is the presenter effectively using interactional resources, e.g. eye-contact, audibility and voice control?)	A	B	C	D
3.3	<b>Visual Resources</b> (is the presentation readable? Is s/he using an effective combination of verbal and visual resources?)	A	B	C	D
4 COMMENTS					
<hr/> <hr/>					

## Acknowledgements

I am very grateful for the comments and feedback received from Christopher Williams, Marina Bondi and the anonymous reviewer that helped me shape my work in its present form.

## References

- Alderson J. C. & B. North (eds.) 1991. *Language Testing in the 1990s: The Communicative Legacy*. New York: Macmillan.
- Baldry A. P. & P. J. Thibault 2006. *Multimodal Transcription and Text Analysis*. London: Equinox.
- Boud D. & N. Falchikov (eds.) 2007. *Rethinking Assessment in Higher Education: Learning for the Longer Term*. London: Routledge.
- Brown A. L. & J. C. Campione 1994. Guided discovery in a community of learners. In K. McGilly (ed.), *Classroom Lessons: Integrating Cognitive Theory and Classroom Practice*. Cambridge, MA: MIT Press/Bradford Books: 229-270.
- Callahan E. 2007. *Effects of Peer Assessment on Attitude Toward Science in High School Students*. PhD Dissertation, Walden University.
- Cohen P. A., J. A. Kulik & C. L. C. Kulik 1982. Educational outcomes of tutoring: a meta-analysis of findings. *American Educational Research Journal* 19: 237-248.
- Douglas D. 2000. *Assessing Languages for Specific Purposes*. Cambridge: Cambridge University Press.
- Dudley-Evans T. & M. J. St. John 1998. *Developments in English for Specific Purposes*. Cambridge: Cambridge University Press.
- Falchikov N. 2001. *Learning Together. Peer Tutoring in Higher Education*. London/New York: Routledge.
- Falchikov N. & J. Goldfinch 2000. Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks. *Review of Educational Research* 70/3: 287-322.
- Frankland S. (ed.) 2007. *Enhancing Teaching and Learning Through Assessment*. Dordrecht: Springer.
- Greenwood C. 1997. Classwide peer tutoring. *Behavior and Social Issues* 7/1: 53-57.
- Halliday M. A. K. 1978. *Language as Social Semiotic. The Social Interpretation of Language and Meaning*. London: Arnold.
- Halliday M. A. K. & C. M. I. M. Matthiessen 2004. *An Introduction to Functional Grammar*. Third Edition. London: Arnold.
- Hughes A. 2003 [1989]. *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- Hutchinson T. & A. Waters 1987. *English for Specific Purposes: A Learning-centred Approach*. Cambridge: Cambridge University Press.
- Jewitt C. & G. Kress 2003. *Multimodal Literacy*. New York: Peter Lang Publishing.
- Jin X. H. 2012. A comparative study of effectiveness of peer assessment of individuals' contributions to group projects in undergraduate construction management core units. *Assessment & Evaluation in Higher Education* 37/5: 577-589.
- Johnson D. W. & R. T. Johnson 1987. *Learning Together and Alone: Cooperative, Competitive, and Individualistic Learning*. Englewood Cliffs, NJ: Prentice Hall.

- Kaufman J. H. & C. D. Schunn 2011. Students' perceptions about peer assessment for writing: their origin and impact on revision work. *Instructional Science* 39/3: 387-406.
- Kress G. & T. van Leeuwen 2001. *Multimodal Discourse. The Modes and Media of Contemporary Communication*. London: Arnold.
- Kress G. & T. van Leeuwen 2006. *Reading Images. The Grammar of Visual Design*. Routledge: London/New York. Second Edition.
- Liu N. F. & D. R. Carless 2006. Peer feedback: the learning element of peer assessment. *Teaching in Higher Education* 11/3: 279-290.
- Malinowski B. 1923. The problem of meaning in primitive languages. In C. K. Ogden & I. Richards (eds.), *The Meaning of Meaning*. London: Routledge and Kegan Paul: 296-336.
- McConlogue T. 2012. But is it fair? Developing students' understanding of grading complex written work through peer assessment. *Assessment & Evaluation in Higher Education* 37/1: 113-123.
- O'Donnell A. M. & K. Topping 1998. Peers assessing peers: possibilities and problems. In K. Topping & S. Ehly (eds.), *Peer Assisted Learning*. Mahwah, NJ: Erlbaum: 255-278.
- Pond K., R. Ul-Haq & W. Wade 1995. Peer review: a precursor to peer assessment. *Innovations in Education and Training International* 32/4: 314-323.
- Sindoni M. G., M. Cambria & F. Stagno d'Alcontres 2007. Test di accreditamento: specifiche elaborate dal CLAM. In G. Di Martino & M. Gotti (eds.), *Sperimentazione e Didattica nei Centri Linguistici di Ateneo*. Napoli: Arte Tipografica Editrice 1: 305-325.
- Sindoni M. G. 2009. Testing English at university. Some reflections on validation. In M. G. Sindoni (ed.), *Procedure di Valutazione nei Centri Linguistici Universitari*, Quaderni di Ricerca del Centro Linguistico dell'Ateneo Messinese 2. Soveria Mannelli: Rubbettino: 139-151.
- Sindoni M. G. & R. Rizzo 2008. Learning and assessment procedures at the University of Messina Language Center (Italy). An integrated approach. In *The 2nd International Multi-conference on Society, Cybernetics and Informatics*. Orlando 29/06-02/07/2008. Orlando: IIIS, Vol. 2, 254-258.
- Sindoni M. G. & R. Rizzo 2009. Learning and assessment procedures at the University of Messina Language Center (Italy). An integrated approach. *Journal of Education, Informatics and Cybernetics* 1/3: 48-52.
- Sindoni M. G. & M. Cambria 2010. Valutare la lingua inglese all'Università: validazione e standardizzazione. In N. Vasta & N. Komninos (eds.), *Il Testing Linguistico*. Udine: Forum: 73-87.
- Stiggins R. J. 1994. *Student-centered Classroom Assessment*. New York/Toronto: Merrill.
- Surowiecki J. 2004. *The Wisdom of Crowds*. New York: Anchor.
- Topping K. 1988. *The Peer-Tutoring Handbook. Promoting Co-operative Learning*. London/Sydney: Croom Helm.
- Tratnik A. 2008. Key issues in testing English for Specific Purposes. *Scripta Manent* 4/1: 3-13.
- Weir C. J. 2005. *Language Testing and Validation. An Evidence-Based Approach*. New York: Palgrave Macmillan.